SAM Hints and Tips

Adam Lyon
Collaboration Meeting Tutorials
February 11, 2003

1. About Dataset Definitions (Sam commands, snapshots, good & evil queries)
2. Problems? No problem! (Checks you do, If things are slow, Getting help)
3. Miscellaneous Tips (Pick_events, SamRoot)
4. SAM Shifts
5. Future Plans
6. SAM usage

Who am I?

◆ I'm not a super-SAM expert

◆ But, I'm involved in tools development with SAM (so I go to the SAM meetings)

◆ I know what the SAM team is worried about

◆ As a sam-admin for clued0 and cab, I'm aware of sam issues

◆ I won't give a general introduction to SAM
  ❖ See `http://d0db.fnal.gov/sam`
  ❖ See `http://d0server1.fnal.gov/www/Computing/Web/documentation.htm`

◆ I'll instead cover some issues that have come up recently

◆ "Hopefully I'll tell you a little more than what you already know; certainly a little more than what I already know."

Where's the talk?

◆ http://www-d0.fnal.gov/~lyon/samTips.ppt

◆ http://www-d0.fnal.gov/~lyon/samTips.pdf

# What is SAM? (more than just green eggs and ham)

◆ **S**equential data **A**ccess via **M**eta-data (Huh?)

◆ Handles file storage
  ❖ Sam keeps track of the location of files (tape, disk)

◆ Handles file meta-data
  ❖ Don't have to know the file name to get files
  ❖ Keeps track of file relationships (parentage; essential for luminosity)

◆ Delivers files
  ❖ You don't care where the files come from

◆ Bookkeeping aid
  ❖ Sam keeps track of your projects
  ❖ Years from now, you can check what files your project analyzed and if there were problems

SAM in the world

- ◆ **SAM is truly global**
  - ❖ Over 20 stations!
  - ❖ Several Fermilab stations
    - central-analysis (d0mino)
    - Clued0
    - CAB
    - Production Farms

Dzero SAM Deployment Map

★ Processing Center
● Analysis site
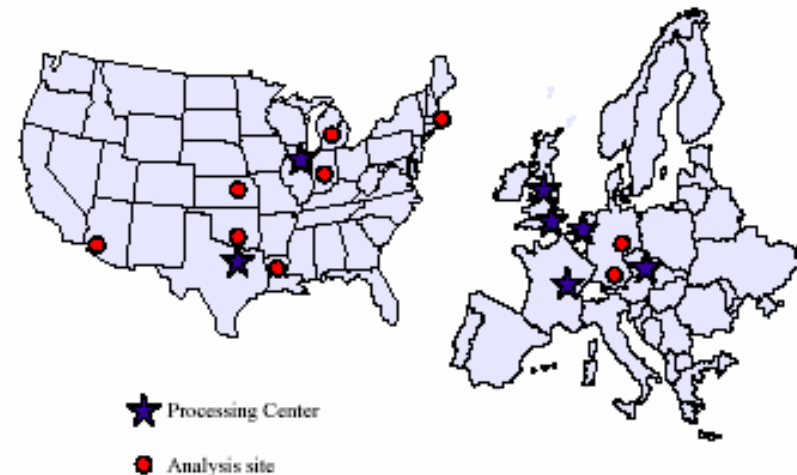
- ◆ **SAM is multi-species**
  - ❖ CDF is becoming SAM-ready
  - ❖ Eventually, SAM will be a Fermilab Product for use by new experiments
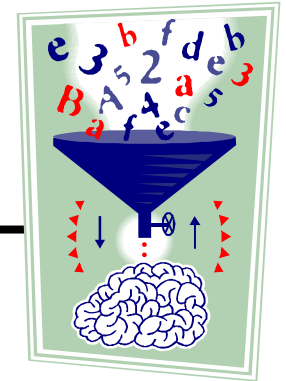
Part 1: About Dataset Definitions

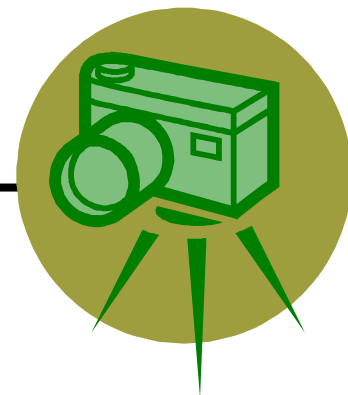**SAM Dataset Definition Editor**

`sam translate constraints`

`sam create dataset`

Creating a dataset definition

◆ Dataset definitions specify the files you want SAM to deliver

◆ Create via the web with the DDE or on the command line

◆ Command line:

❖ Test your query with

```
sam translate constraints
```

❖ Make your definition with

```
sam define dataset
sam create dataset
```

Sam datasets and snapshots

◆ The ***dataset definition*** is the "query"

  ❖ Defines what files you want

◆ The *snapshot* is the list of files that corresponds to the dataset definition

  ❖ The files that correspond to a dataset can change over time
  ❖ A snapshot is a file list for a certain point in time

```
sam define dataset
```
  Defines the dataset but does not make a snapshot

```
sam create dataset
```
  Defines the dataset and makes the first snapshot

Sam datasets and snapshots

◆ `sam submit` (used by d0tools) by default makes a new snapshot
  ❖ Can specify `--dataset-version=last` to use the most recent previous snapshot
  ❖ Useful for reproducing results

◆ `sam run project` does not specify a snapshot. See your python script: `snapshot_version = "new"`
  ❖ If it is set to *last* and you don't have a snapshot, you'll get an error

◆ `sam verify project snapshot --defname=`
  ❖ Shows you what has changed since your last snapshot
    • What files would be added to the snapshot (*e.g.* new files)
    • What files would be removed from the snapshot (*e.g.* bad tapes)

Specifying the definition

◆ `sam translate constraints`

❖ With no options gives you help

◆ Constraints:

❖ Simple query words (see command help)

```
sam translate constraints --runnum=168573
                          --datatier=thumbnail
```

◆ Dimensions:

❖ More complex query language

❖ `sam translate constraints --dim=help`

❖ `sam translate constraints --dim=help --type=datafile`

Dimensions: Use of __set__

◆ To get list of files for a dataset from
   `sam translate constraint –dim="…"`

   ❖ To see the files in your snapshot (what you ran over last) do…
      `dataset_def_name "my_data_set_name"`

      • This just retrieves the list of files in the snapshot
      • It's FAST!
      • But you can't apply additional constraints

   ❖ To see the most recent file set…
      `__set__ "my_data_set_name"`

      • This retrieves the constraints set for the dataset and runs it
      • Can be VERY slow, especially with MC (since lots of MC parameters)
      • SAM people are working to make this better!

Dimensions: Processing unconsumed files (Marco)

◆ Create a "base" dataset

```
sam create dataset --defname="tmb1306"
    --dim="version p13.06 and data_tier=thumbnail"
```

◆ Create the "catch-up" dataset

```
sam create dataset --defname="tmb1306uncomsumed"
    --dim="__set__ tmp1306 minus
            (project-name lyon-tmb1306-% and
             consumed_status consumed and
             consumer lyon)
```

◆ Your projects should be named as "lyon-tmb1306-200302171342"

Dimensions: Removing special runs

◆ Problem: You want to make a dataset definition but exclude special runs

◆ Solution:
   Use `onl_run_type` dimension.

   `--dim='run_number 166780-166789 and`
   `      data_tier raw and onl_run_type Physics'`

   Valid values are `Calibration`, `Cosmics`, `Physics`, `Special`, `Test`.
   Note dimension queries are *case-sensitive* (use the capitalization above)

◆ Confusion: There is also a `run_type` dimension.
   Valid values are: (`calibration, detector, detector data,`
   `monte carlo, online archive, other, physics data taking,`
   `run1, vlpc characterization`).
   Use to separate collider data from MC. Note all lowercase.

Cutting on detector quality

◆ Problem: You want runs where the muon system had good quality

◆ Solution: Use `run_qual_group` and `run_quality` dimensions together

```
--dim='run_number 166780-166789 and data_tier raw and
(run_qual_group MUO and run_quality REASONABLE)'
```

◆ Unfortunately, you can't do multiple groups (but this is coming very soon)

DB query tips

◆ Avoid querying on the file name when you can use meta-data (use
`data_tier, physical_datastream_name, appl_name, run_number, …`)

◆ If you must solely query on the file name, **never** put % first (e.g. %168573%)

  ❖ You lose all indexing – DB will have to scan EVERY entry (millions).

  ❖ Will be horribly slow and you will bog down DB server

  ❖ If you find you can't avoid this usage, send mail to sam-admin so experts can come up with changes to the meta-data.

Filename % badness

◆ Slow DB server response related to users doing %filename query
   ❖ Most queries take a few seconds
   ❖ Examples:
      • '%0000164640%root' hogged DB server for 1.5 minutes
      • '%mcp10%nikhep%' hogged the DB server for 17 minutes (and found no files)
      • 'WZskim-muStream-%p13.05.00' and '%20021221-171855%' Not a DB hogger, but could have been done better

   ❖ Hogging the DB server means that everyone else using that DB server waits until the query is done! We do have several DB servers (command line/web/CAB) but can't make too many more. An idle DB server takes resources

   ❖ Adding checks in DB servers to prevent some long queries

◆ Use meta-data!

Testing your job

◆ If you are testing your job, use a dataset with one file

    ❖ To get the name of a file in your dataset

```
sam translate contraints
            --dim='dataset_def_name mydef'
```

    ❖ Create a new dataset with just that file

◆ Many stations prefetch files (especially on CAB). You are wasting resources if you are just testing your analysis code.

Part 2: Problems?
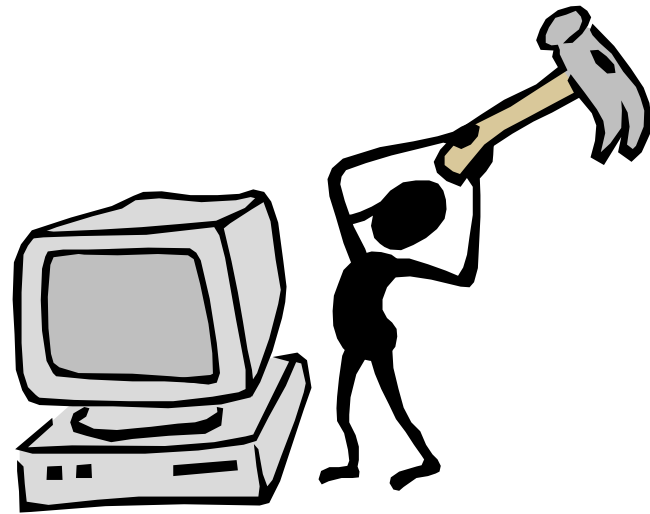
Sam Stability

- ◆ SAM use has increased recently:
  - ❖ New machines:
    - • Farm reached design size
    - • Cab opened to public
  - ❖ Many groups are producing official skims
  - ❖ New MC request system (many parameters, heavy DB load)

- ◆ Complexity of system:
  - ❖ Many problems are fixed in new station versions, but to maintain stability not all stations upgrade
  - ❖ There are lots of active station

- ◆ SAM continues to be stable, though the extra load has caused some glitches

Something goes wrong – what to check

◆ If you see `CORBA.COMM_FAILURE` in the error dump…

　❖ Minor communications glitch

　❖ Relax; Wait a few minutes; Try again

　❖ If no success after several attempts, ask for help (saying how many times you've tried with no success)

◆ On clued0 station:

　❖ Are you running in batch? [Always run in batch]

　　• Interactive use skirts use limits (two simultaneous file transfers will kill a clued0 node)

　❖ Is station set correctly? (for `sam run project`)

　❖ Are you running from a nfs mounted directory (use `/work/chutzpah-clued0/lyon/sam`, not `/disk2/work/lyon/sam`)

A. Lyon (FNAL/CD/D0) – 2003

Sam delivery timeouts
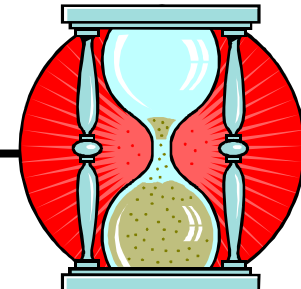
◆ Your job dies with:

```
---------------------------------------------------
%ERLOG-e SAM: PROJECT MASTER:
        Callback error caught in SAMManager::selectForCB()!
        Error message: Timed out waiting for the project master callback!
        File delivery may be slow or unavailable!
         mgr  7-Feb-2003 22:24:47  ReadEvent:read
---------------------------------------------------
```

◆ For some reason, SAM file delivery is slow:
   ❖ The database is busy
   ❖ Station is very busy
   ❖ Station is confused

Make sure your `SAMMangager.rcp` file has

```
    int FileRequestTimeout = 240

    int ProjectMasterTimeout = 240
```

If your project appears to hang...

◆ Check your project state…

◆ Try (for clued0)

```
setup sam

sam dump project --project=(your project)
```

◆ Try (for CAB) [on d0mino]

```
setup sam -q cab

sam dump project --project=(your project)
```

◆ Note this command uses resources
  (don't run it every 50 ms with a script!)

Parsing the output (1)

```
<d0mino> sam dump project --project=tuple1304.mu105

*** BEGIN GPM DUMP ***
Input files: 2310545..2315763   (Your file codes)
  2310545: WZskim-muStream-20021218-122647.raw_p13.04.00, size=0K, unbuffered yet
  2314943: WZskim-muStream-20021219-090757.raw_p13.04.00, size=0K, swapped out,
                           node = d0cs012.fnal.gov, delivered on 10 Feb 13:51:23
  2315038: WZskim-muStream-20021219-184615.raw_p13.04.00, size=0K, unbuffered yet
  2315106: WZskim-muStream-20021220-010232.raw_p13.04.00, size=0K, buffered,
                           node = d0cs012.fnal.gov, delivered on 10 Feb 13:51:23
  2315763: WZskim-muStream-20021220-151308.raw_p13.04.00, size=0K, unbuffered yet
Cached (not buffered) files: (none)    (Not sure why this is none -- should be 3 cached)
Buffered files:
File# 2315106: Delivery time: 10 Feb 13:51:23, overall state: opennow
  Open/Closed by 1/0 consumers, first at 10 Feb 15:13:42
  Consumer states (cid, cfs): (106320,opennow) (You've opened this file)
External files with delivery problems: (none)

Umer contexts (name, state, join time, nSeen): ("Umer" is short for Consumer)
  106320: mverzocc(wz_analyze:v01-00-01), active, 10 Feb 13:51:22, 1 (You've done 1)
Proc contexts (ID: name, state, join time [, current|last]):
  1426433: mverzocc(wz_analyze:v01-00-01)@d0cs012[framework process], busy since
    10 Feb 15:13:42, 10 Feb 13:51:22, 2315106 (You're processing this file)
Processes waiting for call back:
  CID=106320: (none)  (You're job is not waiting because it's actually running)
```

Parsing the output (2)

◆ Waiting for a file delivery

```
Umer contexts (name, state, join time, nSeen):
  106364: jonas(d0reco:t00.92.00), active, 10 Feb 16:24:59, 8
Proc contexts (ID: name, state, join time [, current|last]):
  1426736: jonas(d0reco:t00.92.00)@aldan-clued0[framework process],
   wait since 10 Feb 16:37:43, 10 Feb 16:25:00, 2418188
Processes waiting for call back:
  CID=106364: 1426736@aldan-clued0.fnal.gov:-30151 (10 Feb 16:37:43)
```

◆ All done

```
Input files: 2384432..2384432
  2384432: recoT_all_0000168389_mrg_100-130.raw_p13.05.00, size=0K,
   swapped out, node = aldan-clued0.fnal.gov, delivered on 10 Feb
   13:56:12


Umer contexts (name, state, join time, nSeen):
  106321: clement(d0reco:t00.92.00), finished, 10 Feb 13:53:46, 1
Proc contexts (ID: name, state, join time [, current|last]):
  1426438: clement(d0reco:t00.92.00)@aldan-clued0[framework process],
   finished since 10 Feb 14:39:51, 10 Feb 13:53:46, 2384432
Processes waiting for call back:
  CID=106321: (none)
```

SAM status for Clued0/CAB

◆ Some basic SAM project status at
  `http://www-clued0.fnal.gov/clued0/sam/web/samHealth.php`
  (on clued0) (mine)

◆ Just a prototype right now, but you can check your project and stager status

◆ Note that "ghosts" of old projects pollute the page

Is SAM healthy?
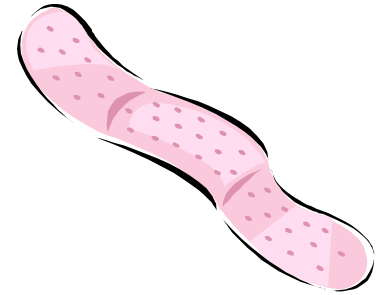
◆ A quick check is:

```
> setup sam
> time sam locate foo
No such file: foo
1
```

◆ Timing on
  ❖ clued0 should be around 3s user, 0.15s system
  ❖ d0mino is longer (8s user, 5s system)
  ❖ If the system time >> user time, something could be wrong with the system

◆ If d0mino is unhappy, Clued0 (uses d0mino cache) and CAB (station runs on d0mino) will also be unhappy

What to do when response is slow

◆ Problem: sam commands take a long time to respond

◆ This may be a d0mino system problem. Check the d0mino performance metrics.

http://d0om.fnal.gov/d0admin/sysperf/
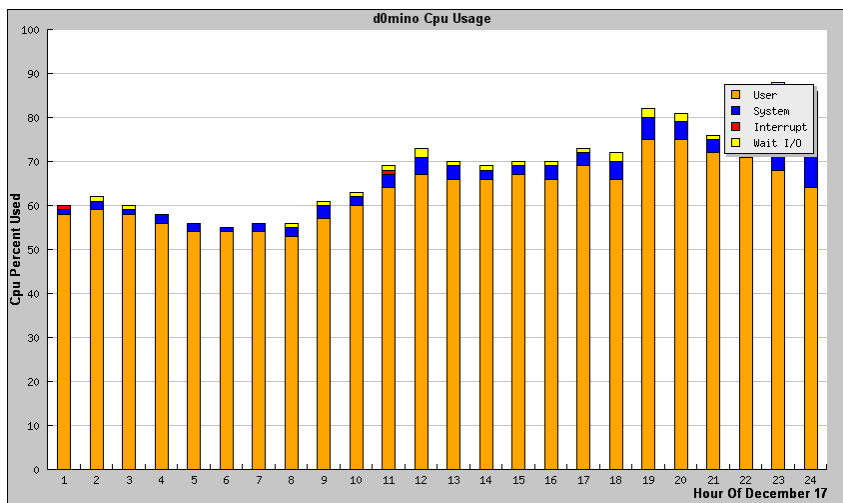
A good day (12/17)
Utilization is now <50%
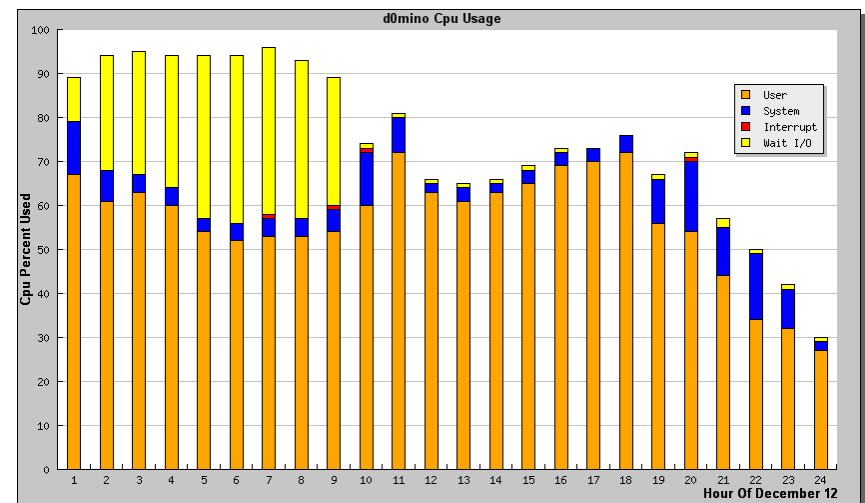
A very bad day (12/12)
A root job was writing to a full disk. Irix got confused and d0mino became bogged down. (Root/Irix problem)

Slow response continued

◆ Problem: Misweb/DDE web sites are slow

◆ Usually caused by usage of d0ora1 (production oracle DB server). **Everybody** talks to d0ora1.

◆ Note that response will be slow early morning (FNAL time) due to backups and other jobs.
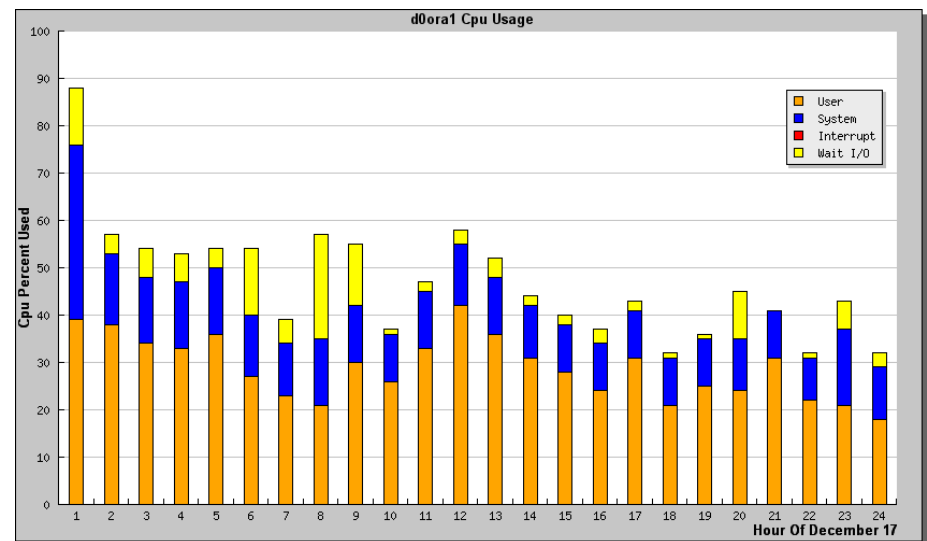
◆ If CPU is maxed then DB response will certainly be slow. But DB can be slow for other reasons too.

When all else fails: Ask for help

◆ CAB: `d0cab-users@fnal.gov`
(`sam-admin@fnal.gov` too)

◆ Clued0: `sam-clued0@fnal.gov`
(we'll send mail on to `sam-admin` if necessary)

◆ Central-analysis: `sam-admin@fnal.gov`

◆ Other stations: your local sam-admin

How to phrase your question

◆ At the very least, give the following:

❖ Your station (clued0, cab, central-analysis,…)

❖ Your project name(s)

❖ Detailed description of the problem (why do you think something is wrong)

◆ Bad:

```
Sam is broken. Please fix it. Hurry up.
```

◆ Good:

```
My project (lyon_124883) on clued0 station has
not received any files for the last three hours.
My job is running on argus-clued0. I have some
log files <here>. Is there something wrong?
You'll get chocolate cake if you fix it quickly!
```

How to stop your SAM project

◆ Kill your batch job

　❖ [Clued0]: `qdel 111222`   (your job #)

　❖ [CAB]:   `qdel 222333@d0mino` (must include @d0mino)

◆ And (please!) stop the project

　❖ `setup sam`

　❖ `sam stop project --station=xxx --project=yyy`

Part 3: Miscellaneous Tips

Pick events

◆ Right now, if you use p13 pick_events is slow

◆ Use p11 instead; but there are some files that p11 will crash

◆ If you get a dspack error, use p13 for those files

◆ The p13 problem is being looked at by experts (but low priority)

SamRoot   (Gabriele Garzoglio)

- Instead of copying root files to your area with SAM, run SAM within root (Sam-Root Serial Interface)
- See `http://d0db.fnal.gov/sam/doc/userdocs/SamRoot.html`

- Use *sam_client_api* A user interface to SAM from C++/root

- Caveats:
  - ❖ Can only read files, not store
  - ❖ *Now works on linux!!!!*

- Like runProject.py, but it runs within root!

- Are people using this interface (We've gotten little feedback)? Please give it a try. Send experiences (good & bad) to d0-atg.

- Should be very useful when we can read thumbnails directly into root

Part 4: Sam Shifts

SAM shifts

- ◆ **What we'd like:**
  - ❖ A system that's up 24/7
  - ❖ And when it's not, monitoring and experts on call 24/7.

- ◆ **What we've got:**
  - ❖ A quite reliable system but with weaknesses (d0mino, d0ora1)
  - ❖ Sam shifters that are not super-experts (but they're learning)
  - ❖ Super-Sam experts that are busy developing new SAM features (and they need sleep too)
  - ❖ Very little coverage during the day Fermilab time (so the busy experts are even busier)
  - ❖ SAM needs you!

Sam response to problems

◆ In the past, response by SAM shifters to problems (e-mail to sam-admin) has been less than stellar – SAM team is trying to improve situation

◆ Kin Yip (BNL) is now SAM shift coordinator

◆ Kin and Wyatt track mail to sam-admin and responses

◆ Better shifter training

◆ The shifter should respond to every message to sam-admin

SAM incident tracking

WEEK	# prob	Fielded shifter	Solved shifter	Fielded expert	Solved expert	Fielded None	Pending	# crashes CAS/CAB/CLU
11/25	23	11	11	9	9	1	2	0/0/0
12/2	22	2	1	19	16	1	4	1/0/0
12/9	30	6	4	23	19	1	5	1/1/0
12/16	23	4	3	16	16	2	0	0/1/0
12/23 – 1/6	14	4	3	10	8	0	3	0/0/0
1/6	18	6	2	11	12	1	2	0/0/0
1/13	22	6	4	15	15	0	1	3/0/0
1/20	25	6	3	19	20	0	2	0/1/3

◆ Situation is improving

◆ SAM shifts may evolve into offline shifts

Part 5: Future Plans

CDF + DØ Togetherness

◆ CDF will use SAM for their file delivery system

◆ CDF has requested changes to the SAM database
  ❖ Luminosity information
  ❖ Links to handle split events (CDF can split parts of an event to different runs)
  ❖ Handle filesets

◆ Since the DB schema is changing, we have some changes of our own:
  ❖ Reduce overloading of concepts in the database (less faked data)
  ❖ Add some features

New Sam Schema
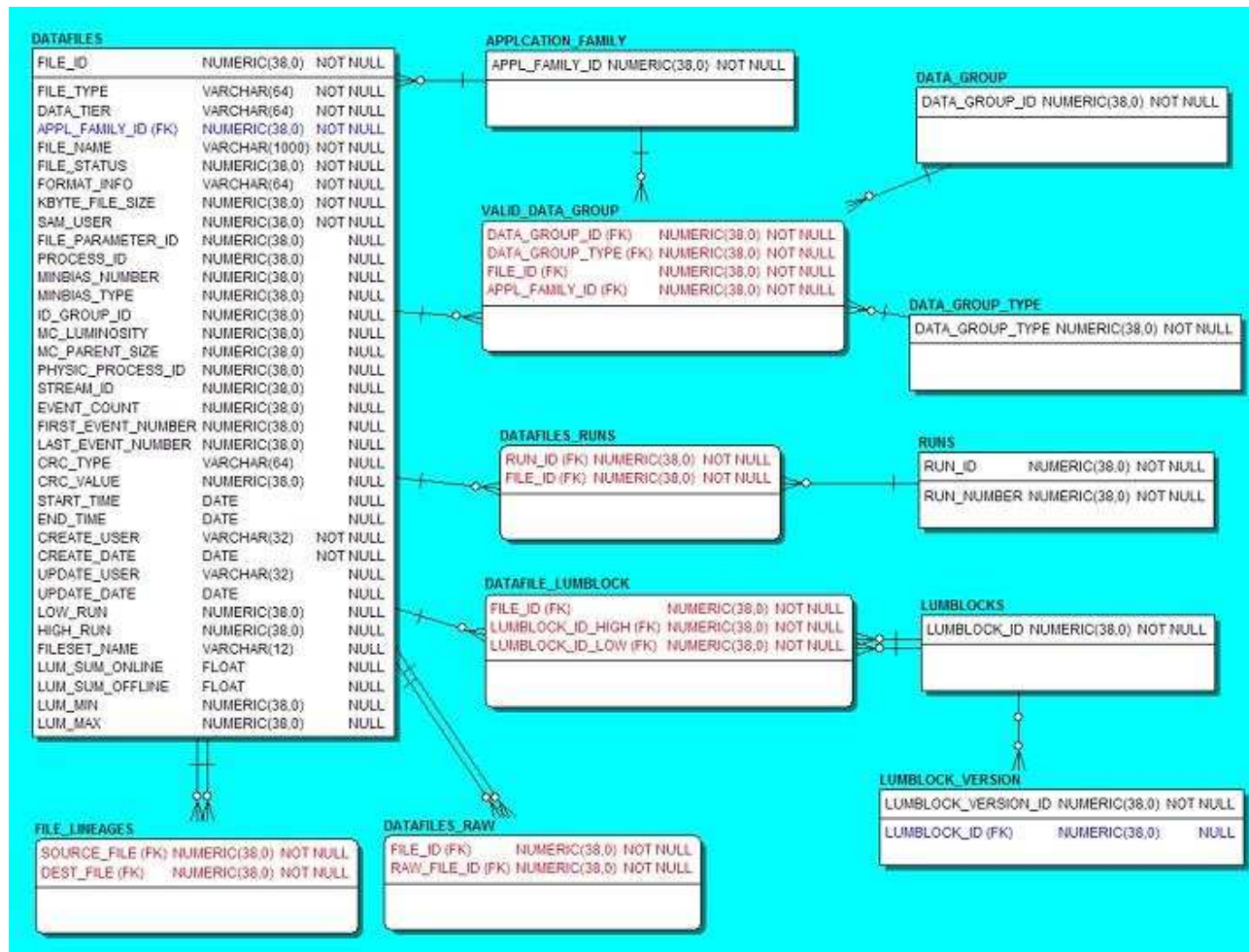
- ◆ Explicit file types (easier queries)
  - ❖ Collider data
  - ❖ MC
  - ❖ Official
  - ❖ User

- ◆ Data groups
  - ❖ e.g. p13Moriond

- ◆ Direct hop to Raw
  - ❖ Easier for lum

Future user features

◆ Working on tools for SAM project archeology

◆ Error handling for `runProject.py`
  ❖ Your job copies files and your disk fills up: SAM thinks you got all requested files

  ❖ For now, you should be keeping track of copy or other system command errors

◆ SAM Project Manager error handling
  ❖ Will no longer throw an exception; your job will end gracefully if SAM error

Part 6: SAM Usage

Gbytes Consumed per Hour on All Stations
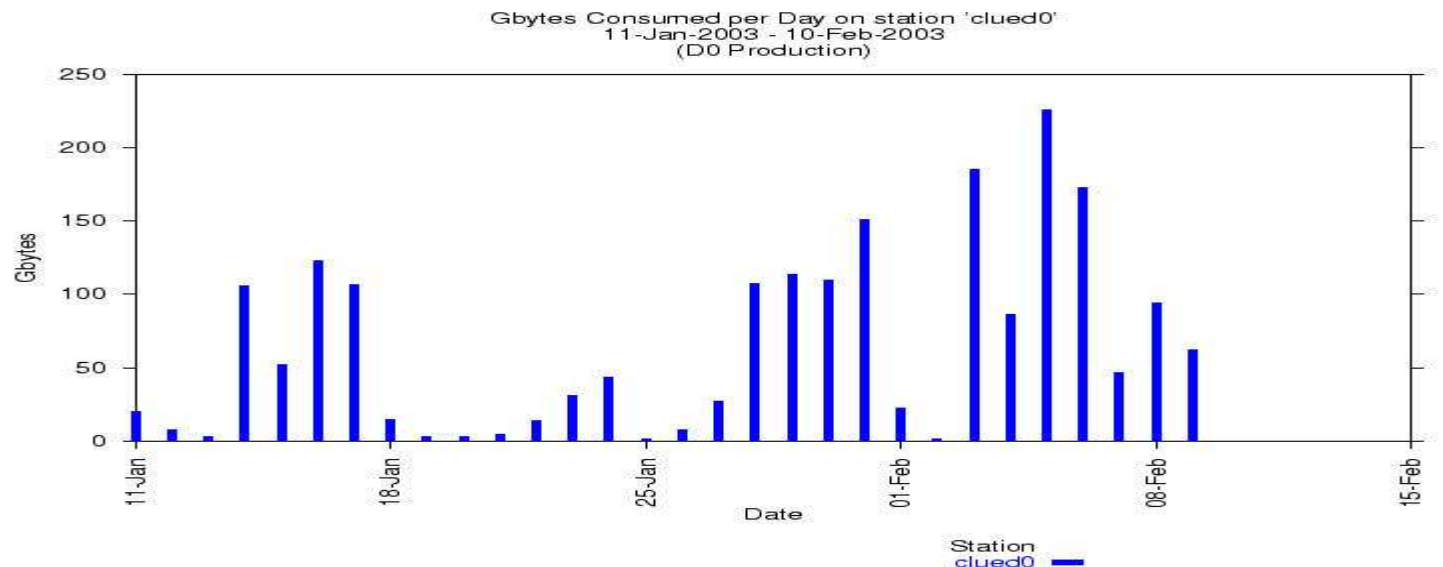09-Feb-2003
(D0 Production)
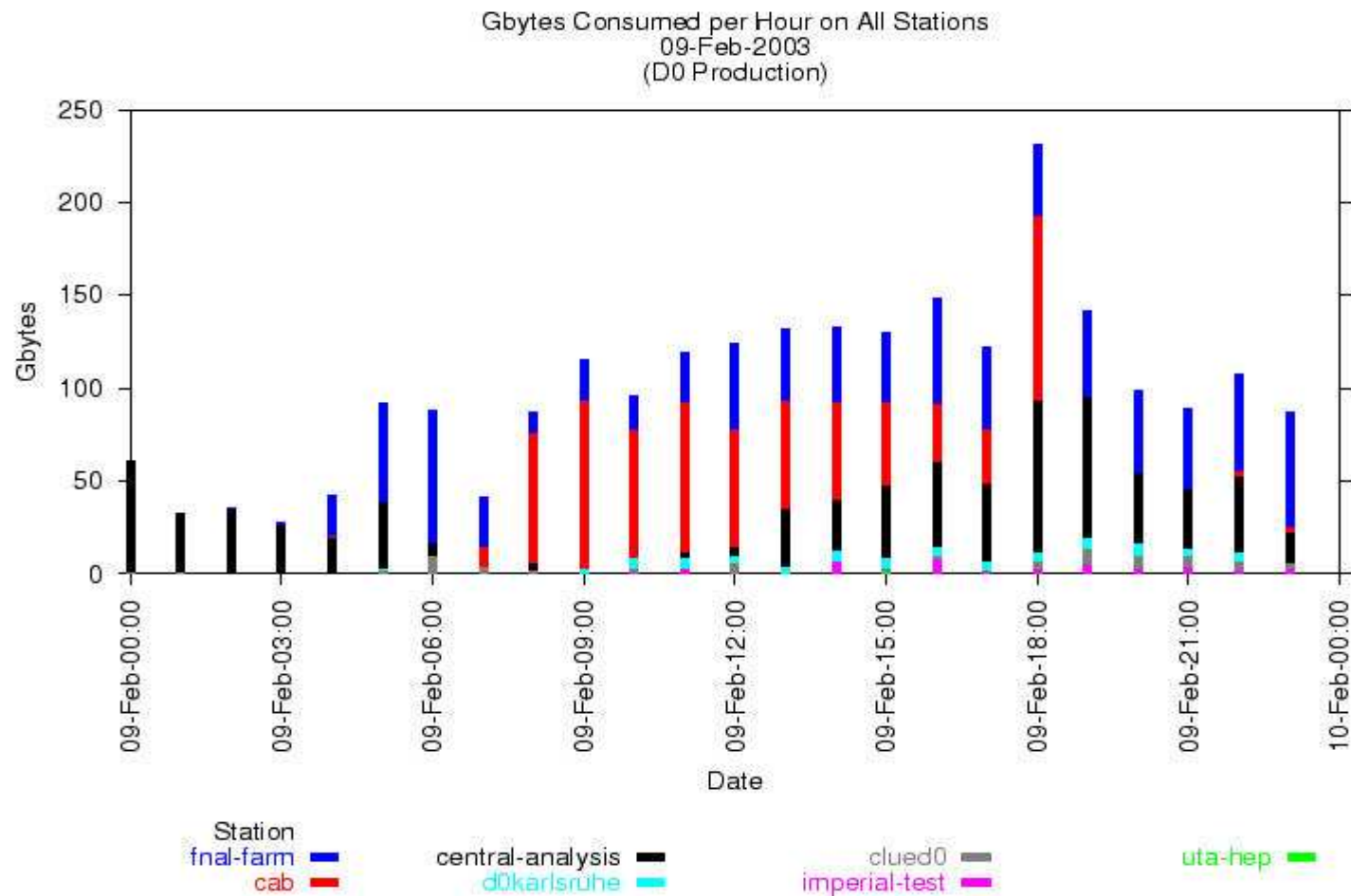
Sam on Clued0
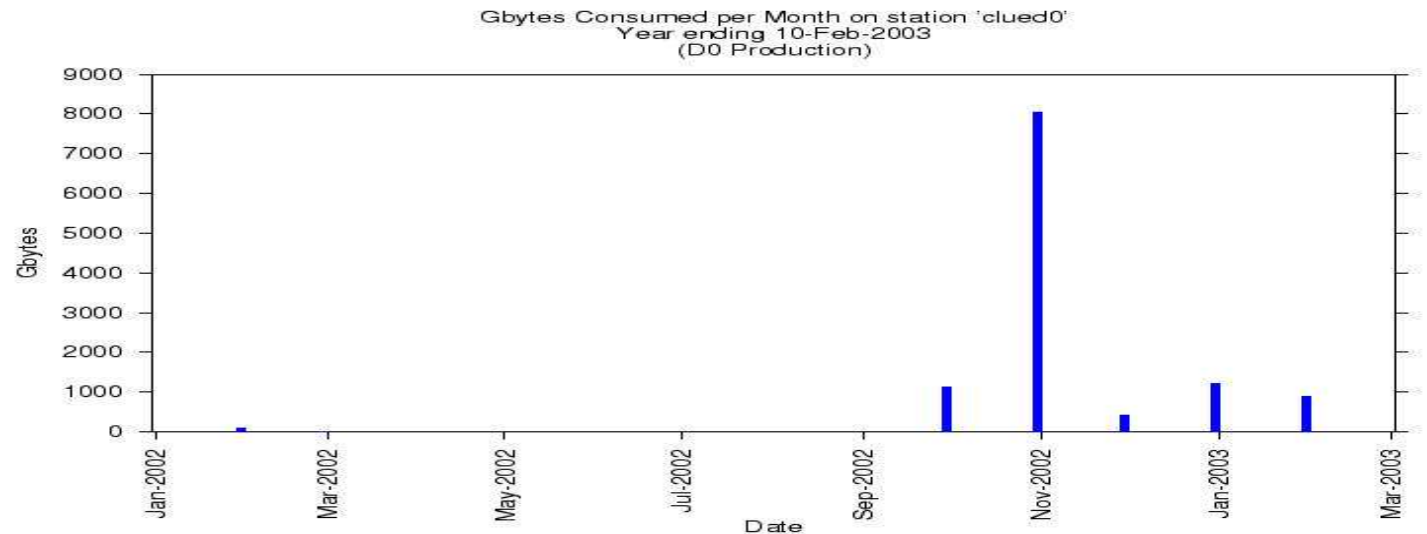
◆ Installing/Configuring/Testing SAM on clued0 has been a long process. Thanks to Sinisa, Lukas, Dugan, …

◆ Clued0 officially opened for SAM business on December 6, 2002
  ❖ 40 worker nodes (so max 40 SAM jobs)
  ❖ SAM only available through batch system
  ❖ Limited to 6 simultaneous downloads from enstore/d0mino (< 1TB/day)
  ❖ No stagers in DAB due to network hubs
  ❖ Clued0 SAM-admins are Adam Lyon & Jon Hays
  ❖ If problems, send mail to sam-clued0@fnal.gov. We triage and will forward to sam-admin if we're stumped.

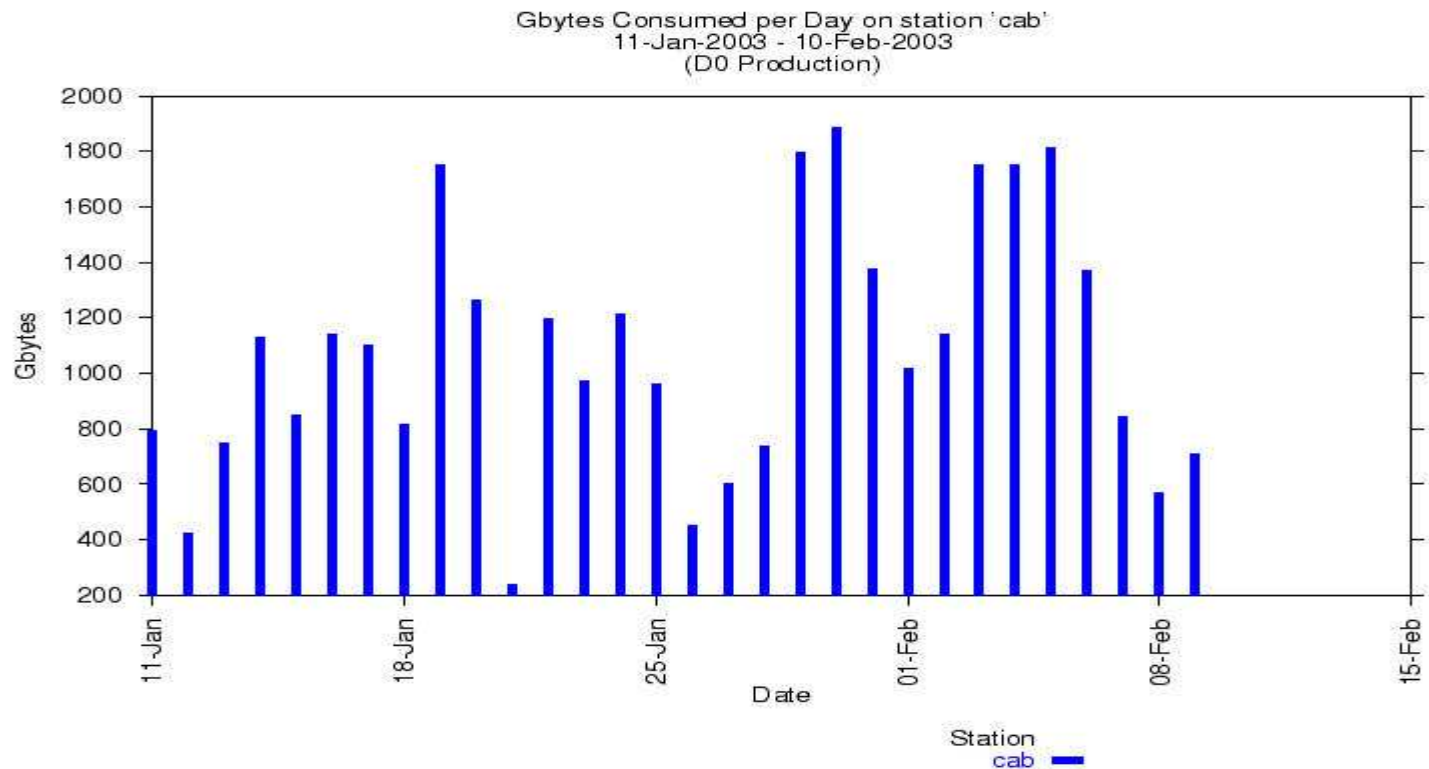# Clued0 Sam Usage

◆ From 12/6 to present:

❖ >2TB consumed

◆ Saving files back to SAM works
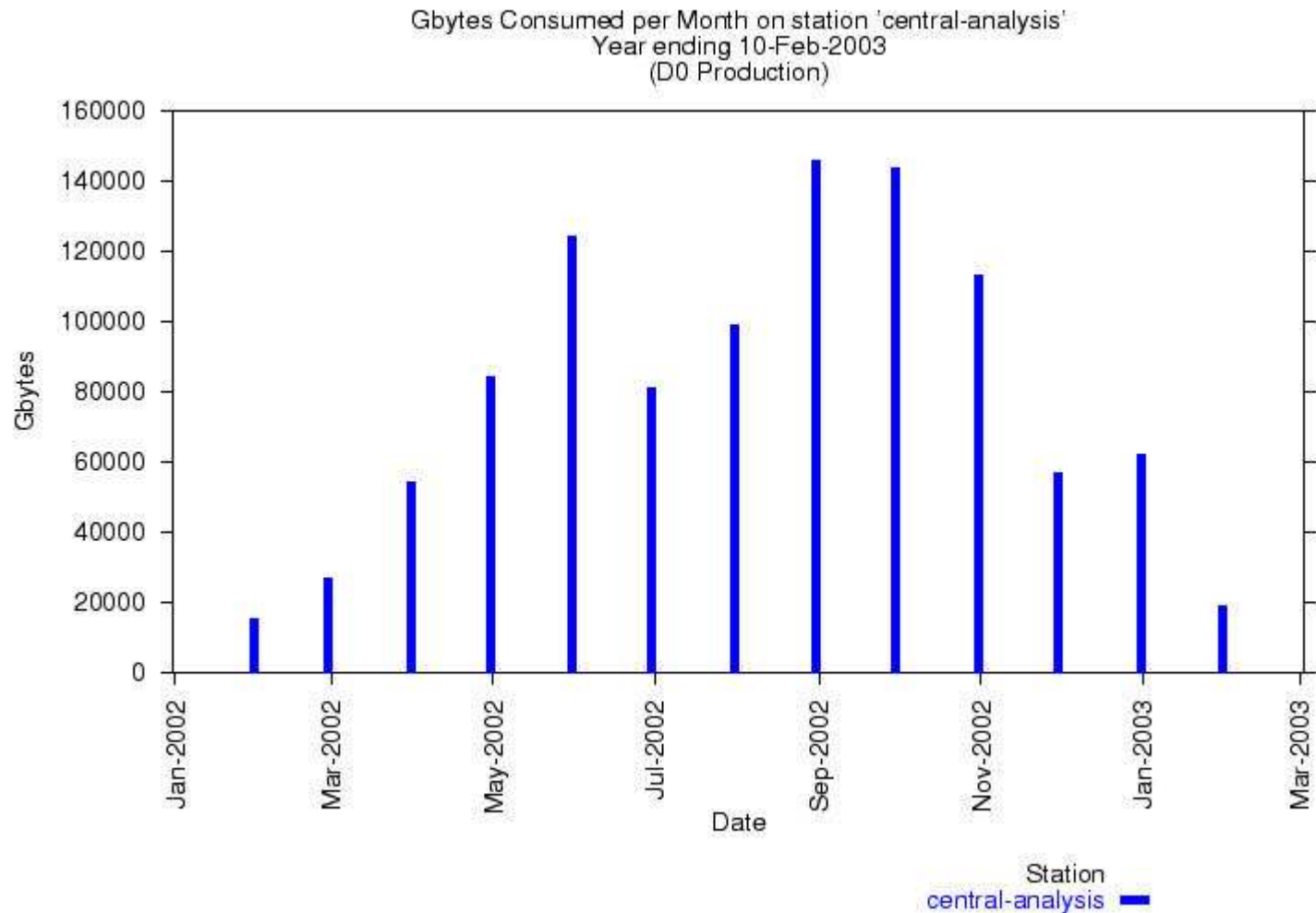
◆ But we're still small potatoes

A. Lyon (FNAL/CD/D0

Gbytes Consumed per Month on station 'clued0'
Year ending 10-Feb-2003
(D0 Production)

Gbytes Consumed per Day on station 'clued0'
11-Jan-2003 - 10-Feb-2003
(D0 Production)

Sam on CAB

- ◆ ~160 worker nodes; designed for SAM use
- ◆ Station and cache are on d0mino
- ◆ Opened for business on 1/29

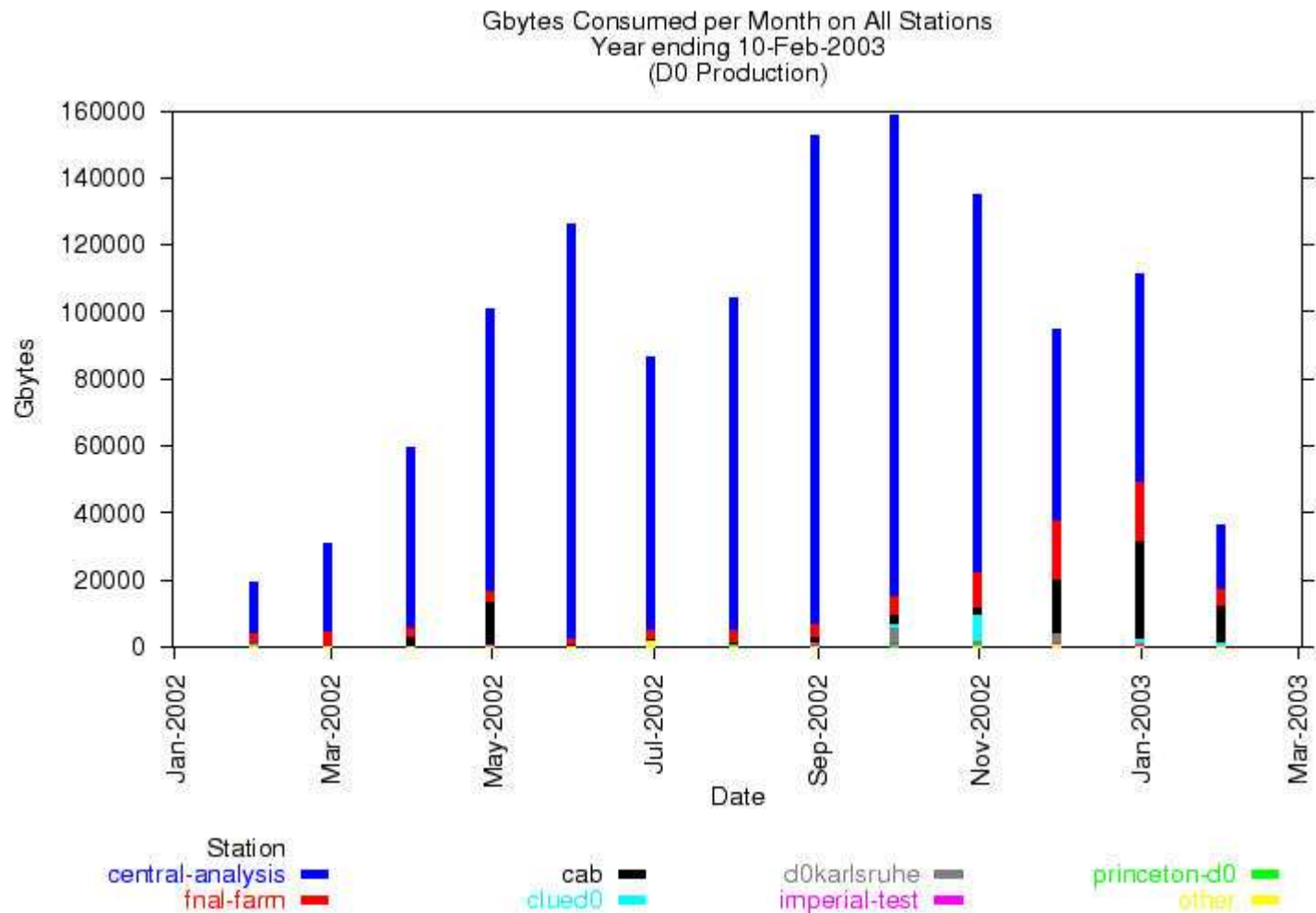Gbytes Consumed per Day on station 'cab'
11-Jan-2003 - 10-Feb-2003
(D0 Production)

Sam on central-analysis (d0mino)

Gbytes Consumed per Month on station 'central-analysis'
Year ending 10-Feb-2003
(D0 Production)

All of SAM

Gbytes Consumed per Month on All Stations
Year ending 10-Feb-2003
(D0 Production)

Summary

- ◆ Create your datasets carefully -- use meta data as much as possible

- ◆ SAM is quite reliable, but things can go wrong
  - ❖ Ask for help when you need it, but realize that SAM is a complicated system and expert coverage is not 24/7

- ◆ You can contribute to SAM's up time by becoming a SAM shifter

- ◆ SAM DB changes will make your queries more efficient and more analysis friendly

- ◆ SAM is used around DØ, Fermilab and the World!